Session 2: The Production of Science as a Tournament



Figure 1 Growth of Tournament Theory in the Literature.

1)Tournament races with priority as incentives

Knowledge is a public good that is not used up when one person consumes it. That you learn that $a^2 + b^2 = c^2$ does not use up the fact nor removes it for the next person to learn and use. Once knowledge is publicly available, **no one with sufficient understanding** can be excluded from using it. For knowledge that has positive uses, the social optimum is to make it available to all at zero price.

But there is little incentive for you to produce knowledge if you have no control over it and cannot extract some rewards for your effort. The result is too little incentive to produce knowledge for public use.

Merton's solution: "I propose the seeming paradox that in science, private property is established by having its substance freely given to others who might want to make use of it." By publishing in reputable peer review journals the scientist claim ownership for the idea, which creates a reputation among peers, and leads to rewards from society. "Since positive recognition by peers is the basic form of extrinsic reward in science, *all other extrinsic rewards, such as monetary income from science-connected activities, advancement in the hierarchy of scientists ... derive from it.* ... this kind of extrinsic reward system provides great incentives for engaging in the often arduous and tedious labors required to produce results that enlist the attentions of qualified peers and are put to use by some of them. (Merton 1988, p. 621) – Hmm economists might put greater emphasis on money reward.

Citations are critical in this process: a citation "registers in the enduring archives the intellectual property of the acknowledged source by providing a pellet of peer recognition of the knowledge claim, accepted or expressly rejected, that was made the source. (Merton, 1988, p. 622). In short, "The institutions of science have evolved to motivate scientists to contribute freely to the common wealth of knowledge ..., just as they can freely take from that common wealth what they need. Moreover, since a fund of knowledge is not diminished through exceedingly intensive use by members of the scientific collectivity that virtually free and common good is not subject to the tragedy of the commons."

What Merton did not stress was that the incentive of gaining credit for an idea/result generates a **"tournament"** in which scientists race to be the first to publish an idea/finding, with positive and negative consequences for scientific activity.

"When I was a young faculty member struggling to earn tenure, I was denied authorship on a paper that represented a major scientific advance in my field. It is an injustice that I have since learned is pervasive throughout the scientific community, stemming from a hierarchy based on seniority alone. Failing to credit the junior scientists who make many of our original discoveries not only undermines the importance of this younger class of researchers, but actually threatens scientific progress. Their paper, the first to be published on the association of actin and the bacterium, has been cited 765 times. Coming second, my paper has received ... 233 citations.. Twenty years later, I still don't trust confiding new findings to other researchers. My takeaway lesson was that the safest strategy was to divulge my results only after they were accepted for publication. All's not fair in Science and Publishing" (Southwick, https://www.the-scientist.com/critic-at-large/alls-not-fair-in-science-and-publishing-40801).

Assignment: Search the Internet, Harvard's Hollis journals to find Southwick's paper and the paper that he claims usurped his work and its senior authors. There are other clues in the article. If you need, co-citations might help are co-citations. You find other papers that cited many of the same articles on notion should be similar to given paper. PS <u>https://ufhealth.org/news/2010/southwick-named-2010-harvard-university-advanced-leadership-fellow</u>

Famous Tournament Races

1) The Double Helix aka Honest Jim not being so honest

"We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure hasnovel features which are of considerable biological interest ... It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material"







Battles in the Race for Priority on DNA

1) Wilkins refused a request by Pauling to see some diffraction patterns

2) The National Foundation for Infantile Paralysis (founded by FDR!) "March of Dimes" turned down Watson's request to switch the subject of his fellowship to DNA. Promising a funding agency to do one thing and then doing something different is a well-established scientific tradition.

3) Watson attributed Pauling's successes with proteins to building models with incomplete data, then comparing them to see which model came close to accounting for the X-ray patterns. He based his analysis on this strategy.

4) Sir Lawrence Bragg told Watson and Crick to stop working after their first DNA model failed, but they proceeded to do what they wanted.

5) Watson and Crick never told Pauling they found a fatal error in his model, because they feared that he would try again and might beat them out.

6) Franklin provided the critical X-ray data (unbeknownst to her) that insisted the bases be placed at the center of the structure and the sugar-phosphate chains outside, but the theory paper did not reference to her or Wilkins.

7) Crystallographer Jerry Donohue's advice was critical in getting the bases right, but not given any authorship.

8)Watson and Crick tell Wilkins and Franklin immediately since the model's "proof" is their data.

Questions: What is the key to Watson and Crick getting result first – the entrepreneurial acts? the single-minded pursuit of a goal? Watson says it was his concentration on the genetics that was critical.

How important is the putting together a model vs finding the evidence?

Would more names appear on papers today for far less contribution than Franklin and Donohue made to the

Watson-Crick paper? How would you test the hypothesis that today papers spread credit more broadly?
of authors on paper has risen --- because papers require more technology and thus need more experts.
Were there more acknowledgments on papers a half century or so ago?

Is the Watson-Crick breakthrough two great minds cracking a problem or the result of COMMUNITY activity (though part of the community did not know it was contributing) so that someone else would have discovered the double helix shortly?

How might we predict when a particular question will be answered? What would you infer from many teams working on the question than on other questions?

QUANTUM NANOTECH MACHINES "whose movement obey the weird rules of quantum mechanics, which state that an object can absorb energy only in discrete "quanta" and be in two places at once" (Science, Jan 29, 2010) (http://www.sciencemag.org/content/327/5965/516.full.pdf) – SCIENCE **Breakthrough of the year**

In 2009-2010, five groups came within a few dozen quanta of creating a genuine nanotech machine, and one

may have reached it. The competitors: teams of condensed matter physicists, experts in optics and astrophysics. "The field has exploded in terms of the number of people working in it ... bringing toolboxes from different fields" trying to make a machine that has a half-quantum of energy and jiggle with zero-point motion that oscillates at a well defined frequency and absorbs energy only in quanta proportional to frequency. They seek to remove all but the last half-quantum by lowering oscillator's temperature nearly to absolute zero. To make energy quanta as large as possible, they etched beams of semiconductor that vibrated at high frequencies—up to 1 billion cycles per sec. To make the beams as cold as possible, they stick them in liquid-helium refrigerators thousandths of a degree above absolute zero. The beam's motion causes the voltage to vary. Then experts in quantum optics began experimenting with laser light and mirrors. Each photon must absorb a quantum of energy from the mirror to make up for the energy it is lacking. So that "detuned" light saps energy from the oscillator.

The competitors: Aspelmeyer University of Vienna team used a mirror on a beam;

Kippenberg Swiss Federal Institute of Technology, Lausanne shined light into a glass ring that served as both optical cavity and oscillator.

Hailin Wang team at the University of Oregon, Eugene used a glass bead in a similar way.

Keith Schwab of **Caltech** used a silicon-nitride beam 30 micrometers long and about 150 nanometers thick and wide that thrums at 6.3 megahertz to get down to 4 quanta

THE WINNER: In the race to the ground state, the straightforward approach may have won out. Andrew Cleland, John Martinis, and colleagues at the University of California, **Santa Barbara** used a "brute force" combination of passive cooling and a very high oscillator frequency Kippenberg argues: "It's not a purely mechanical oscillator"

THE NEXT PHASE: Oskar Painter, **Caltech.** "You've got phonons, photons, and electrons... That's where the revolution is going to come from." convert optical signals to microwaves and vice versa "Systems either behave quantum mechanically or classically," says Nergis Mavalvala, **MIT** "Is there something murky in between?

Question: What has happened to careers of winners vs other competitors? What is the current state of work

USPTO has over 10,000 patents on nanotechnology. How many have produced product sold on market. Etimes "Global nanotechnology business is expected to be around \$125 billion by 2024, with startups to large companies, including GE, Intel, Samsung, BASF and AkzoNobel, developing products and solutions around nanotechnology."

THE LATEST: Meet the nanomachines that could drive a medical revolution

https://theconversation.com/meet-the-nanomachines-that-could-drive-a-medical-revolution-58107 A single-atom heat engine Roßnagel et al Science 15 Apr 2016: Vol. 352, Issue 6283, pp. 325-329

Questions: Does the race for priority speed up the production of knowledge? Did the different techniques illuminate different aspects of the issue? Will the straightforward cooling or the laser light method yield more science in the long run? Which will "scale up" to provide something useful? What proportion of scientific activities are parts of races? How does the Roffnagel et al single atom machine connect to the 2010 tournament?

Different Prizes in Sciences

Nobel Prize: Every year in Sweden and Norway, deserving men and women are honored in chemistry, economics, literature, peace, physics and physiology or medicine. Each receives a diploma, a gold medal and more than \$1 million if there is only one winner in that category. Two or three people being honored in a single category split the funds. Alfred Nobel, a Swedish inventor, used nearly 95 percent of his fortune to fund the Nobel Prize.

Lasker Awards: awarded by the Lasker Foundation since 1945, are given to people "who have made major advances in the understanding, diagnosis, treatment, cure and prevention of human disease." Its categories include the Albert Lasker Basic Medical Research Award, Lasker-Bloomberg Public Service Award, Lasker-DeBakey Clinical Medical Research Award and Lasker-Koshland Special Achievement Award in Medical Science. Winners awarded \$250,000, and nearly 100 people who have received these honors have gone on to receive a Nobel Prize.

Wolf Prize: Awarded by the Wolf Foundation since 1978, with \$100,000. The categories include agriculture, physics, medicine, chemistry, mathematics and arts. A number of people feel that winning one of these awards is the second-most prestigious honor behind the Nobel Prize. In fact, about half of those who win Wolf Foundation honors in physics have gone on to claim a Nobel Prize. Also, there is no Nobel Prize in the category of agriculture, and the Wolf Prize is the most prestigious honor you can earn in that field.

Breakthrough Prize:Dating to 2012, it has already earned quite a reputation due to the amount of money its winners receive – \$3 million – and its televised awards ceremony, which are designed to mimic the Oscars. Nearly \$22 million in prizes were claimed at the 2015 ceremony, which took place in the heart of Silicon Valley, in Mountain View, Calif. Breakthrough Prizes are awarded in fundamental physics, life sciences, and mathematics.

Kavli Science prizes for advances in astrophysics, nanoscience and neuroscience. Consisting of a scroll, medal and cash award of one million dollars, a prize in each of these areas is awarded every two years beginning in 2008.

Mathematics Prizes: Fields medal, for people under 40, just \$15,000. Other awards are the Wolf Prize of the Wolf Foundation of Israel, the Leroy P. Steele Prize of the American Mathematical Society, the Bôcher Prize, Cole Prizes in algebra and number theory, and the Delbert Ray Fulkerson Prize, all presented by the American Mathematical Society. The Norwegian Academy of Science and Letters awards the **\$985,000 Abel prize**. The Clay Mathematics Institute of Cambridge, Massachusetts (CMI) has named seven "Millennium Prize Problems," selected by focusing on important classic questions in mathematics that have resisted solution over the years with a\$1 million allocated to each. A cash prize of \$100000 has been offered for proof/counterexample to Beal's conjecture.

Hypothesis: Since math requires less \$\$ resources, should get more DC recipients, wider distribution of winners by country. Easy to check.

AAAS Awards To recognize scientists, journalists, and public servants for significant contributions to science and to the public's understanding of science, the Association gives awards at the AAAS Annual Meeting immediately following the award year: AAAS Philip Hauge Abelson Prize; AAAS Award for International Scientific Cooperation; AAAS Award for Public Understanding of Science and Technology; AAAS Mentor Awards ; AAAS Scientific Freedom and Responsibility Award; AAAS Kavli Science

The Royal Society Prizes for Science Books decided by the Royal Society, the UK national academy of science.;; "the Booker Prize of science writing". Many national prizes are only open to citizens of the countries concerned. Australia has prize; NZ has prize ...

2)Tournaments in Scientific Careers

Tournaments are the major incentive mechanism for hiring and promoting people. A tournament ranks people by their performance and gives those who come at the top promotions and prizes. Tournaments exaggerate differences — the person who gets the paper to the journal a week before a second person gains most of the credit.

For executive tournament see Ian Gregory-Smith and Peter W. Wright Winners and losers of corporate tournaments Oxford Economic Papers, Volume 71, Issue 1, January 2019, Pages 250–268, where winner of top job gets promotion/pay and loser leaves or gets salary increase as well to keep at job.

In science/math the person who pioneered/developed a strategy that possibly inevitably lead to a solution gets less credit than the person who completes the proof.

In math fame went Perelman, who "proved" Poincare's Conjecture not to Richard Hamilton, who laid out the strategy of using Ricci flows to solve the problem. The desire for priority led to disagreement about whether Yau's group of Chinese students had filled in enough details in Perelman's proof to claim credit (see Nasar and Gruber New Yorker piece-- https://www.newyorker.com/magazine/2006/08/28/manifold-destiny) because Perelman wrote very concisely.

Are tournaments efficient? The economics model of a tournament gives prizes to persons depending on how they end up in a rank order. The winner gets the biggest prize but the **marginal product of the tournament exceeds the marginal product of the winner** because it includes the output generated by all participants, losers included, which would make tournament potentially super-efficient/highly profitable to extent the prize motivated immense work and that foregone productivity/rewards were comparable to productivity in tournament.

But rank order tournaments run the risk of sabotage/reduction of the value of research: I break into your lab and mess up your experiment/vote against funding ideas that may undo my ideas; Andrew Wiles worked on Fermat's Theorem in his attic while telling no one because he feared someone else might finish the job faster. There is a conflict between rank order and helping others with their papers and reporting results.

Compare tournaments to alternative ways to motivate workers. An ideal reward system is self employment — aka sales contract – in which worker produces output Q and maximizes PQ(e) - C(e) where P is price, e is effort, C is cost of effort. In equilibrium PQ' = C': marginal cost of effort equals marginal value of effort, which we can rewrite

as Q'/C' = 1/P to highlight that ability and taste for work enter commensurately. This is an "ideal contract" with no shirking or conflict (beyond you fighting with yourself).

Piece rates are a form of self-employment: given the piece rate workers differing in ability and/or cost of effort decide how much to produce, which yields the self employment equilibrium. The piece rate would depend on value of product in the competitive market.

Productivity: A has twice the productivity of **B**, **Qa** = 2 Qb so per hour worked A earn twice what **B** earns Cost of effort: Cw (e) = 1/2 C(e) where w is workaholic so cost of effort is half that to other person. These cases are indistinguishable. What matters in effort and hourly pay is Q'/C', which says that if the more able person has greater taste for work/lower cost of effort, pay inequality rises – 4 x more inequality in pay.

Common mode of pay is **time rate** — Workers gets wage W and produces Q(e). Employer keeps worker as long as Q(e) > Qm; fires worker otherwise. Worker maximizes W-C(e) . What is equilibrium? Qm, the minimal effort to maintain the job. Firm will have to monitor worker to make sure worker performs at Qm or better.

In tournament employer chooses how much to pay winner **and losers**. The winner is person with highest score (S) - Score = Effort x Ability + luck = EA + u.

Grant competitions are tournaments.

Fellowship competitions are tournaments.

Up and out tenure systems are tournaments – promote 2 of 10 assistant professors. If criterion is top 2, pure tournament. But piece rate: publish 2 papers in A-journals; 4 in a B rated journals; and 7 elsewhere and you get tenure – can also be tournament if funding limits say cannot hire all those above the cut point. Tournaments may be between persons of comparable ability and between persons with different abilities.

Equal ability tournament: For simplicity consider a tournament with two players, so the the probability person 1 wins is $P = F(S_1-S_2>0)$. If the two people have identical Es and As, luck determines who wins. But even if your EA> my EA, the tournament might motivate me because maybe with luck, my experiment works and I win the prize. Luck is motivator.

Firm chooses prize structure: W for the winner, L for the loser to get the most production at least cost. This depends on WORKER RESPONSES to the incentive; If workers respond with lots of effort to W-L, firm will set large W-L.

Worker chooses effort depending on C(E). If the worker maximizes the expected value of WP + (1-P) L – C(E), differentiating gives (W-L) dP/dE - C' = 0 where dP/dE = effect of effort on the chance of winning. The optimal level of effort equates the marginal increase in the chance of winning x extra reward from winning to the marginal cost of effort. Thus, Bigger W-L --> more effort; Bigger dP/dE --> more effort.

With **homogeneous workers** having same ability and effort, the chance of winning depends on u: P = prob (EA+ $u_1 > \text{EA}+u_2$) => prob ($u_1 > u_2$) – LUCK, where the distribution of luck f(u_1 - u_2) has mean 0 and a variance. With two EQUALLY ABLE workers the EXPECTED P is $\frac{1}{2}$. Each worker produces so that (W-L)f(0) = C', where f(0) =1/2 = (W-L)1/2 = C'.

Firm sets W and L to gain most effort from workers. If P is the value of output per unit of effort, the firm maximizes P(E + E) - (W+L) = 2PE - (W+L) subject to supply constraint that (W+L)/2 = C(E) — ie that the firm pays enough for worker's cost of effort to keep them at the workplace.

Substitute the constraint into the equation and the maximand becomes 2PE-2C(E) = 2 (PE-C(E)).

To obtain W and L differentiate the maximum wrt to each: (P-C')dE/dW = 0 and (P-C')dE/dL = 0.

The equilibrium condition is C' = P. This is the same first best solution as in the self-employment and piece rate cases. It equates the marginal value of the output of effort, P, with the marginal cost of effort C'. Now go back to the workers decision that said that they produce so that (W-L)f(0) = C'. Thus, the firm sets W-L = P/f(0).

If workers are risk averse they will put out less than "first best" effort if W-L is large so firm will find it optimal to offer lower W-L contracts.

Note the difference between tournament and piece rate pay. If you produce for a piece rate with some disturbance, only your disturbance affects your outcome. In a tournament, the disturbances of the other competitors affects your outcome, creating a collective problem. If everyone has a bad day or shirks in a tournament, the winner still gets W whereas if everyone does badly with a piece rate everyone gets paid less.

Unequal Ability Tournament: When workers have different abilities (Moldovanu & Sela, AMERICAN ECONOMIC REVIEWVOL. 91, NO. 3, JUNE 2001(pp. 542-558) optimal is to give multiple prizes, because better to give marginal incentive to everyone than only to the potential #1s. With workers having different abilities the chance of winning depends on the ability difference and on luck:

 $P = prob (AE_1 + u_1 > AE_2 + u_2) \longrightarrow prob (AE_1 - AE_2 > u_1 - u_2)$

If one person knows the other is better and the random component is small, will not have a genuine tournament. Think of playing jeopardy against WATSON. You will not review your knowledge of trivia since you are sure to lose.

But can have handicaps — affirmative action; veterans preferences; etc. In an unequal field, we need to do something to create competition. In horse racing handicap is adding weights to saddles to even out the competition -->more unpredictable outcomes \rightarrow more money for the track. We expect firms or market to sort workers into what Lazear and Rosen "Rank-Order Tournaments as Optimum Labor Contracts." Journal of Political Economy, Vol. 89, No. 5, (October 1981), pp. 841-864. call major/minor league tournaments to have people of similar ability.

For review of many studies see B.L. Connelly L Tihanyi T. Russell Crook K. Ashley Ganglof Tournament Theory: Thirty Years of Contests and CompetitionsJournal of Management Vol. 40 No. 1, January 2014 16–47 https://www.researchgate.net/publication/275441821_Tournament_Theory_Thirty_Years_of_Contests_and_Competitions/ ink/574c666808aec98852669c77/download.

Construct	Description	Pertinent Citations
Tournament	A contest wherein actors compete for a prize that is awarded based on relative rank.	Lazear & Rosen, 1981 Rosen, 1986
Prize	Reward for tournament winner(s) designed to incent the effort of all contestants. This may be monetary	Knoeber & Thurman, 1994
	or have monetary value attached to it, but it could also be about prestige, privilege, or the possibility of competing in successive tournaments.	Moldovanu, Sela, & Shi, 2007
Prize spread (for "prize," some also use compensation, wage, pay, or inter-rank; for "spread," some also use dispersion, disparity, inequality, or gap)	In sequential tournaments, the difference between the prize for winning the current tournament and that for winning the next highest level. In promotion contests this reduces to the wage spread between workers at their current level and what they would be earning at the next level.	Becker & Huselid, 1992 Messersmith, Guthrie, Ji, & Lee, 2011
Prize optimization	The prize spread that maximizes the ratio of actor effort to prize. If too small, actors are not incented to exert effort. If too high, actors take on additional risk of losing the contest and need to be separately compensated for that risk.	DeVaro, 2006 Kepes, Delery, & Gupta, 2009
Tournament size	The combination of a tournament's breadth (i.e., number of unique competitors) and depth (i.e., number of possible levels).	Che & Gale, 2003 Boudreau, Lacetera, & Lakhani, 2011
Win percentage	The likelihood that any given actor will win a prize. This is an important predictor of motivation.	Chen, Ham, & Lim, 2011 Taylor & Treadon, 2002
Actor heterogeneity	Differences among actors that could influence tournament variables and their final relative rank.	Bothner, Kang, & Stuart, 2007 Shaw & Gupta, 2007
Handicapping	Adjusting incentives or processes to account for heterogeneity, increasing (decreasing) the win percentage for disadvantaged (advantaged) actors.	Pfeifer, 2011 Frick, 2003
Tracking	Creating subcontests (i.e., tracks within contests) to account for heterogeneity so actors may compete with a more homogeneous subgroup.	Gomez-Mejia, Trevino, & Mixon, 2009 Nippa, 2010
Sequential elimination (hierarchical tournaments)	Actors that win a tournament then compete in another tournament against other winning actors.	Choi & Gulati, 2004 O'Neill & O'Reilly, 2010

Table 1 Primary Constructs

Grade curves as example of a tournament. Consider the output of students facing the following grade curves:

- 1. Everyone in the class gets a B regardless of how they do on the exam. This is a "socialist" solution.
- 2. Top student gets an A and the rest get Fs. This is maximum inequality
- 3. There are 10% As, 40% Bs, 40% Cs, 10% Fs

Output in the class is measured by the sum of correct answers from the group. We expect an inverse U-Curve with an optimum level of inequality. I <I* and I> I* reduces output

Freeman-Gelber (2010) used a maze experiment that finds this pattern with one striking addition. Not only do the number of mazes solved differ with the incentives but so too does the pattern of cheating on the mazes.



Misreporting mazes to experimenter

Output

Treatment	Mean (SD) Mazes Solved		Mean (SD) Mazes					
			Misreported					
Full Information	(1) Round 1	(2) Round 2	(3) Round 1	(4) Round 2				
Equal prize	12.65	15.55	.09	.18				
	(4.86)	(7.08)	(.33)	(.58)				
Multiple prize	12.40	18.56	.13	.51				
	(4.44)	(6.87)	(.54)	(1.44)				
Single prize	11.74	16.10	.19	.46				
	(5.24)	(7.88)	(1.20)	(1.15)				
No Information								
Equal prize	12.46	12.88	.09	.13				
	(4.68)	(5.62)	(.33)	(.49)				
Multiple prize	11.91	16.36	.17	.32				
	(5.27)	(7.89)	(.84)	(1.62)				
Single prize	12.31	16.53	.12	.23				
	(5.05)	(7 38)	(39)	(68)				

Inequality (conditional that incomes sum to output)

Source: Tabulated from the experiment described in text. The sample size is 78 subjects for each

of the six treatments shown, and 468 subjects total.

Table 4 Poisson regression coefficients and standard errors for determinants of total number ofmazes misreported in Round 2

	Full Information			No Information			
	(1) All	(2) Top	(3) Bottom	(4) All	(5) Top	(6) Bottom	
	Subjects	Half	Half	Subjects	Half	Half	
EQUAL							
MULTIPLE	1.01	1.29	.94	21	.33	50	
	(.41)**	(.57)**	(.49)*	(.58)	(.79)	(.88)	
SINGLE	.50	1.75	71	.55	.61	.45	
	(.43)	(.55)***	(.82)	(.44)	(.76)	(.53)	
Mazes	.38	1.87	.46	.74	.85	.72	
Misreported	(.05)***	(.87)**	(.09)***	(.07)***	(.96)	(.11)***	
in Round 1							
Constant	-1.76	-2.56	-1.39	-2.16	-2.75	-1.75	
	(.35)***	(.50)***	(.41)***	(.33)	(.70)***	(.41)***	
# obs	234	127	107	234	127	107	
Log Pseudo-	-177.52	-79.77	-88.16	-102.20	-40.41	-59.31	
Likelihood							

The number of mazes misreported in Round 2 is regressed on dummies for the incentive conditions, the number of mazes misreported in Round 1, and a constant term. Columns 1 and 4 display results for a regression in which the sample is all participants; Columns 2 and 5 display results for a regression in which the sample is all participants with Round 1 ranks 1 to 3 (inclusive); Columns 3 and 6 display results for a regression in which the sample is are gression in which the sample is all participants with Round 1 ranks 1 to 3 (inclusive); Columns 3 and 6 display results for a regression in which the sample is all participants with Round 1 ranks 4 to 6 (inclusive). The sample size is larger for the top half than the bottom half because tied scores were assigned the same rank; results are similar with other ways of assigning ties. Robust standard errors, clustered by group, are in parentheses. *** denotes significance at the 1% level; ** 5%; * 10%.

Table 3 Regression coefficients and standard errors for determinants of change in mazes solved

	(1) Round 1	(2) Round 1	(3) Round 1	(4) Round 1	(5) Round 1	(6) Round 1	(7) Round 1	(8) Round 1
	Top Half,	Top Half,	Bottom	Bottom	Rank = 1,	Rank = 1,	Rank = 6,	Rank = 6,
	No	Full	Half, No	Half, Full	No	Full	No	Full
	Information							
EQUAL					-		-	
	5.70	2.45	1.98	4.22	5.89	3.33	1.67	3.83
MULTIPLE	(1.03)***	(1.31)*	(.87)**	(1.06)***	(2.16)**	(1.49)**	(.92)*	(1.54)**
SINGLE	5.35	2.71	1.95	12	5.75	3.26	4.31	.91
	(1.02)***	(1.42)*	(1.01)*	(.95)	(1.71)***	(1.65)*	(1.29)***	(1.31)
Constant	27	3.34	1.32	2.41	69	3.27	.42	2.09
	(.69)	(.76)***	(.60)**	(.68)***	(1.48)	(.90)***	(.55)	(.96)**
R-Squared	.22	.05	.05	.20	.22	.09	.29	.19
# obs	127	127	107	107	47	47	35	36
								-

from Round 1 to Round 2, by position in Round 1

The dependent variable is the change in the number of mazes solved from Round 1 to Round 2.

EQUAL, MULTIPLE, and SINGLE represent dummies for the equal, multiple, and single prize

conditions, respectively. Columns 1 and 2 display results for regressions including participants 3)Gender, tournaments, ability: why are women relatively underrepresented at the top of some fields? Hypothesis 1: They are risk averse, per Muriel Niederle & Lise Vesterlund, 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?," The Quarterly Journal of Economics, MIT Press, vol. 122(3), pages 1067-1101 The task is to add up sets of five 2-digit numbers. The numbers are randomly drawn and presented in the following way: 21 35 48 29 83 Same distribution of skills for W and M





People decide whether to be paid by piece rate,

Task 1 — Piece Rate: players receive 50 cents per correct answer.

Task 2 — Tournament: player who solves the largest number of correct problems in the group receives \$2 per correct answer, while the others receive no payment

Given choice, women shun tournament : The worst men more likely to choose tournament than best women!





Proportion of Participants Selecting Tournament for Task 3 Conditional on Task-2 Tournament Performance Quartile (Panel A) and Task-3 Choice Performance Quartile (Panel B)

Both men and women guess they are higher ranked than they in fact are, but overconfidence gap bigger for men

TARLE IV

DISTRIBUTION OF GUESSED TOURNAMENT RANK								
	N	We	omen					
	Guessed rank	Incorrect guess	Guessed rank	Incorrect guess				
1: Best	30	22	17	9				
2	5	3	15	10				
3	4	2	6	5				
4: Worst	1	1	2	1				
Total	40	28	40	25				

With guessed rank most of male-female disappears ... so it is male overconfidence



Proportion of Participants Who Select Tournament for Task 4 Conditional on Task-1 Performance Quartile (Panel A) and Guessed Piece-rate Rank (panel B).

Does this contribute to few women at top of SE, law, management- tournament avoidance for risk aversion?

Hypothesis 2: Time Crunch: They Do More Housework Than Men January 19, 2010 By Jill Laster

When the biologist Carol W. Greider received a call from Stockholm last fall telling her she had won a Nobel Prize in Physiology or Medicine, she wasn't working in her lab at the Johns Hopkins University. The professor of molecular biology and genetics was at home, folding laundry. Ms. Greider does many of the household chores, but she isn't alone. A number of her female colleagues also do more around the house than their male partners. "It is not just housework. For women with kids, it is all the other stuff: scheduling sports and play dates, play dates, remembering all of the calendar events for the whole family," said Ms. Greider, who has two school-age children.

Michelle R. Clayman Institute for Gender and Research at Stanford University study found that **female** scientists do 54% of their core household tasks, such as cooking, cleaning, and laundry—about twice as much as their male counterparts. (Paid help and children made up some of the difference.) ... senior and junior faculty members put in similar hours. Women also worked about 56 hours a week, almost same number of hours as men. Men contributed more to home repair, finance, and yard and car care. But those tasks took about one-quarter of the 19.3 hours a week spent in a home on core household tasks. "Some studies of faculty productivity have found that women faculty may produce fewer articles, but the ones they do produce tend to be cited more frequently," Ms. Sheridan said. "But in an academic institution where the number of your publications or grants is the thing that is most highly valued, that is a problem." But Sifan Zhou and Freeman: female papers have fewer citations due in part to homophily in citations and male predominance in science.

Retention after the postdoctoral period: NAS study shows that women made up 18% of the applicants for tenure-track positions in chemistry at Research I institutions between 1999 and 2003, although women earned 32% of the Ph.D.'s in chemistry. In biology, women made up 24% of the applicants for tenure-track positions, although they earned 45% of the Ph.D.'s. One possible solution could be for universities to create more-flexible benefits packages that allow men and women to hire household labor .

Hypothesis 3: Ability at upper tail: Some argue that few women are at the very top of many occupations, especially math-oriented science because women have lower test scores in math and mechanical skills than men (while having much higher scores in reading and writing) and/or because women have lower variance of skills. A

group with a greater variance will have disproportionately many people at the upper and lower tails. If a profession picks upper tail people randomly and more men are at the upper tail, it will be disproportionately male.

What is the evidence on gender differences in distribution of abilities?

Giuso, Monte, Sapienza, Zingales Science 30 May 2008 –PISA tests show women consistently better in reading and men better in most countries in math.

Hedges/Nowell Science, 7 July 2005, vol 269 meta analysis shows differences in means, notes areas where women do better but that variances almost always greater for men. >95% column translates into relative numbers.

Subject area	d		VP	Tail region			
Subject alea			٧n	≤10%	≥90%	≥95%	
Reading comprehensic Project Talent NLS-72 NLSY HS&B NELS:88	n -0.15 -0.05 -0.18 0.002 -0.09	(0.013) (0.027) (0.031) 2 (0.017) (0.020)	1.16 (0.015) 1.03 (0.028) 1.16 (0.036) 1.10 (0.024) 1.16 (0.023)	1.71 (0.088) 1.15 (0.13) 1.50 (0.19) 1.07 (0.099) 1.75 (0.14)	0.90 (0.051) 0.94 (0.11) 0.83 (0.12) 1.03 (0.096) 0.80 (0.072)	1.00 (0.08) 0.81* (0.14 	
Vocabulary Project Talent NLS-72 NLSY HS&B	0.25 -0.06 -0.03 0.07	(0.013) (0.027) (0.031) (0.017)	1.05 (0.014) 1.00 (0.027) 1.08 (0.034) 1.05 (0.023)	0.89 (0.050) 1.02 (0.12) 1.20 (0.15) 0.84 (0.082)	1.57 (0.081) 0.89 (0.11) 0.87 (0.12) 1.06 (0.098)	1.50 (0.01 ⁻ 0.87 (0.15) <u></u> 1.06 (0.14)	
Mathematics Project Talent NLS-72 NLSY: AR MK HS&B NELS:88	0.12 0.24 0.26 0.08 0.22 0.03	(0.013) (0.027) (0.031) (0.031) (0.022) (0.020)	1.20 (0.015) 1.05 (0.028) 1.25 (0.039) 1.19 (0.037) 1.16 (0.026) 1.06 (0.021)	1.00 (0.055) 0.72 (0.090) 1.84 (0.23) 0.99 (0.13) 0.77 (0.078) 0.97 (0.082)	1.33 (0.069) 1.76 (0.019) 1.90 (0.24) 1.70 (0.21) 1.67 (0.14) 1.34 (0.11)	1.50 (0.01 2.34* (0.36 2.20 (0.39) 1.90 (0.34) 2.06 (0.26) 1.64 (0.18)	
Perceptual speed Project Talent: TR CC Ol	-			2.17 (0.11) 1.79 (0.092) 1.50 (0.077)	0.82 (0.047) 0.73 (0.044) 1.00 (0.055)	1.00 (0.08(0.81 (0.06)	
NLS-72 NLSY: CS NO HS&B	-0.23 -0.43 -0.23 -0.21	(0.027) (0.031) (0.031) (0.022)	1.04 (0.028) 0.98 (0.031) 1.08 (0.034) 1.15 (0.025)	1.54 (0.17) 1.60 (0.20) 1.50 (0.19) 1.49 (0.13)	0.70 (0.089) 0.41 (0.077) 0.69 (0.10) 0.73 (0.075)	0.69 (0.12) 0.38 (0.11) 0.67 (0.14) 0.79 (0.12)	
Science Project Talent: PS BS NLSY NELS:88	0.50 0.29 0.38 0.11	(0.013) (0.013) (0.031) (0.020)	1.28 (0.017) 1.15 (0.015) 1.42 (0.044) 1.14 (0.023)	0.57 (0.038) 0.78 (0.046) 0.92 (0.13) 0.87 (0.076)	2.83 (0.15) 2.00 (0.10) 3.40 (0.45) 2.04 (0.16)	7.00 (0.65) 7.20 (1.6) 2.50 (0.28)	
Social studies Project Talent NELS:88	0.31 0.04	(0.013) (0.020)	1.26 (0.016) 1.14 (0.023)	0.89 (0.050) 1.23 (0.10)	2.29 (0.12) 1.59 (0.13)	3.50 (0.27) 1.74 (0.19)	
Nonverbal reasoning Project Talent NLS-72	0.04 -0.22	(0.013) (0.027)	1.04 (0.013) 1.15 (0.031)	1.00 (0.055) 1.49 (0.16)	1.09 (0.059) 0.74 (0.092)	1.00 (0.080 0.67 (0.12)	
Associative memory Project Talent NLS-72 HS&B	-0.32 -0.26 -0.18	(0.013) (0.027) (0.022)	0.82 (0.011) 1.01 (0.027) 1.14 (0.025)	1.56 (0.080) 1.44 (0.16) 1.23 (0.11)	0.50 (0.035) 0.70 (0.089) —	0.43 (0.048 0.69 (0.12) —	
Spatial ability Project Talent HS&B	0.13 0.25	(0.013) (0.022)	1.27 (0.028) 1.27 (0.028)	0.82 (0.047) 0.79 (0.079)	1.86 (0.095) 1.90 (0.17)	2.33 (0.35) 2.39 (0.30)	
Mechanical reasoning Project Talent NLSY	0.83 0.72	(0.012) (0.030)	1.45 (0.019) 1.74 (0.055)	0.36 (0.029) 0.60 (0.094)	8.50 (0.059) 8.00 (1.3)	11.00 (1.2) 10.90 (2.8)	
Electronics information Project Talent NLSY	1.22 0.72	(0.011) (0.030)	2.72 (0.035) 1.56 (0.049)	0.44 (0.033) 0.62 (0.096)	15.20 (1.3) 8.00 (1.3)	∞ 9.90 (2.5)	
Auto and shop informat	tion 1.02	(0.029)	2.34 (0.073)	0.44 (0.079)	66.3 (27)	464 (702)	

* These figures are for the 97th percentile.



Giuso, Monte, Sapienza, Zingales Science 30 May 2008 - PISA tests show women consistently better in reading

Machin and Pekkarinenn Science 28 Nov 2008



Variance ratios in countries that participated in PISA 2003. In the majority of countries, male variability is higher than female variability in both reading and math.







Fig. 1. Theoretical normal distributions for males (orange line) and females (green line) when their means are identical and the M:FVR – 1.2. The schematic on the right shows a blowup of the distributions in the region from 3.8 to 4.2 standard deviations above the mean. Brown, area of overlap of the 2 distributions; green and orange, areas unique to females and males, respectively.

Recent US evidence

The table below provides a breakdown of NAEP score, by subject and gender, in 2013, Consistent with the international PISA results for 15 year olds, there is a very large gender gap in reading at all three NAEP grade levels, and the differences are quite large,

Average NAEP scale scores and gaps by gender, subject and grade, 2013						
		Boys	Girls	Gap		
	Grade 4	219	225	-7		
Reading	Grade 8	263	273	-10		
	Grade 12	284	293	-10		
	Grade 4	242	241	1		
Math	Grade 8	285	284	1		
	Grade 12	155	152	3		

significant at the 0.05 level except grade 8 math. Data from the National Assessment of Educational Progress.

In contrast, the discrepancies in math scores, while statistically significant in all grades except eighth, are generally quite small in magnitude (i.e., the average score for boys is higher than that for girls, but not by much).

New studies show Females show more sustained performance during test-taking than males

Pau Balart & Matthijs Oosterveen Nature communications Nat Commun 10, 3798 (2019). 67-019-11691-y Females tend to perform worse than males on math and science tests, but they perform better on verbal reading tests. Here, by analysing performance during a cognitive test, we provide evidence that females are better able to sustain their performance during a test across all of these topics, including math and science (study 1). This finding suggests that longer cognitive tests decrease the gender gap in math and science. By analysing a datase with multiple tests that vary in test length, we find empirical support for this idea (study 2).



Fig. 1 Performance throughout the test for males and females in Ireland. The figure is based on the PISA 2009 and uses Locally Weighted Scatterplot Smoothing (LOWESS) to visualize the relationship between the probability to answer a question correctly and the position of the question in the test. Source data are provided as a Source Data file (study 1)

THE GENDER GAP CRACKS UNDER PRESSURE: C, Cotton, F McIntyre, J.price NBERWP 16436 Using data from multiple-period math competitions, we show that males outperform females of similar ability during the first period. However, the male advantage is not found in any subsequent period or even after a two-week break from competition. Some evidence suggests that males may actually perform worse than females in later periods. The analysis considers various experimental\treatments and finds that **the existence of gender differences depends crucially on the design of the competition and the task at hand**. Even when the male advantage does exist, it does not persist beyond the initial period of competition.

	Round					
	1	2	3	4	5	
Race Treatment						
Male	0.57	0.49	0.53	0.57	0.61	
	[0.02]	[0.03]	[0.02]	[0.02]	[0.03]	
Female	0.47	0.44	0.52	0.54	0.59	
	[0.03]	[0.02]	[0.03]	[0.02]	[0.03]	
Not-a-Race Treatment						
Male	0.64	0.64	0.66	0.68	0.68	
	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	
Female	0.63	0.64	0.65	0.68	0.67	
	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	
Reading Race Treatment						
Male	0.51	0.53	0.58	0.53	0.52	
	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]	
Female	0.58	0.57	0.65	0.60	0.60	
	[0.04]	[0.04]	[0.03]	[0.04]	[0.04]	
Repeat Treatment						
Male	0.78	0.64	0.68	0.52	0.65	
	[0.03]	[0.03]	[0.04]	[0.04]	[0.03]	
Female	0.76 [0.05]	$0.64 \\ [0.04]$	0.66 [0.05]	0.54 [0.05]	$0.64 \\ [0.03]$	

Table 2: Fraction Correct By Round For All Treatments

Cells report the average fraction correct by round, gender, and treatment. Standard errors are in brackets.

Under Pressure: Gender Differences in Output ... under Competition and Time Constraints

Olga Shurchkovy The Journal of the European Economic Association

One explanation for gender inequality stems from the interaction between competition and two pressure sources, namely, task stereotypes and time constraints. This study uses a laboratory experiment and finds that women under perform the men in a high- pressure math-based tournament, women greatly increase their performance levels and their willingness to compete in a low-pressure verbal environment, such that they actually surpass the men, largely due to the fact that extra time in a verbal competition improves the quality of women's work, reducing their mistake share. On the other hand, men use this extra time to increase only the quantity of work, which results in a greater relative number of mistakes.

RESULT 1. Under high time pressure with a math task: (a) Men and women do not differ in terms of their scores in the piece-rate treatment, but in the tournament, men significantly outperform the women; (b) Men are significantly more likely to self-select into a tournament than women.



Figure 1a. Distribution of Math Scores by Gender Under Piece-Rate, High Time Pressure



Figure 1b. Distribution of Math Scores by Gender Under Tournament, High Time Pressure

But tournaments can induce excessive effort from workers R. Sheremeta Pros and Cons of Workplace Tournaments <u>https://wol.iza.org/uploads/articles/302/pdfs/pros-and-cons-of-workplace-tournaments.pdf</u>

Pros

- Tournaments can create powerful competitive incentives, motivating individuals to exert effort levels well above those predicted by the rational decision-making model.
- Tournaments provide non-monetary incentives in the form of recognition and winning.
- When compared to other compensation schemes, tournaments may require less information about individual performance.
- Common shocks, such as stock market fluctuations, have less of an effect on tournament-based incentives.
- Tournaments play an important function of matching better workers to better jobs.

Cons

- The win-or-lose structure of tournaments creates some winners at the expense of many losers, leading to substantial payoff inequality.
- Relative incentives create a "discouragement effect," causing lower-ability workers to cut back effort or withdraw entirely from competition.
- Workers view each other as competitors when using relative incentives, resulting in more selfish and less helpful behavior.
- Tournaments may encourage counterproductive behaviors such as cheating, sabotage, and collusion.
- Women may be discouraged from participating in tournaments, even when they are more capable and have better skills than men.

Many experiments show efforts "above predicted"



And another experiment where sabotage reduces total output



In science, potential explanation for much "slippery/bad" behavior is tournament stucture; also it offers explanation for high hours worked by grad students/post-docs as they strive for "faculty job" – possibly at cost of total well-being. BUT IT MAY STILL DO WHAT SOCIETY WANTS: producing fastest development of ideas and spread to improve well-being.